

Global Terrorism Analysis

Minal Kharbade ,Deveshya Gupta

EDA CAPSTONE PROJECT

ALMABETTER ,BANGLORE

ABSTRACT:

- Terrorism has become more concentrated in past few years, so we need to improve our security to deal with it .
- We are analyzing the parameters of successful terrorist attack and countries which are affected by the major terrorist attack.
- To strengthen the security system, they can use analyzed data on various parameters.
- This journal gives detail information on feature extraction and the complete data visualization is done in this dataset.
- Global terrorism analysis project involve some steps :
 1. Loading data from dataset.
 2. Cleaning the data.
 3. Extracting the statistics from dataset.
 4. Exploratory data analysis.
 5. Insights from the analysis.
 6. Conclusion

PROBLEM STATEMENT:

- The Global Terrorism Database includes systematic data on domestic as well as international terrorist incidents that have occurred during this time period and now includes more than 180,000 attacks.
- It is important to analyze and aggregate the available data to better understand its various properties.
- My main objective of these project is to help the security system to strengthen their resources by using this analyzed data.

DATA SUMMERY:

- The dataset provided was cleaned some of its column contains null values.
- The dataset contains 5 rows and 135 column but for proper analysis we have taken 5 rows and 17-18 column .
- We have used the following column from the dataset.

1. Year	7. City	13.wounded
2. Day	8.latitude	14.group
3. Month	9.longitude	15.target type
4. Country	10.Attack type	16.weapon type
5. State	11.target	
6. Region	12.killed	

INTRODUCTION:

Terrorism has become more concentrated in past few years , so we need to improve our security to deal with it. We need to analyze the parameters of successful terrorist attack and countries which are affected by the major terrorist attack . Also we can say terrorism is very important factor now days .

We perform the data cleaning over the dataset. further we will analyze the various parameter related to it .We will try to analyze the data using plots , correlation and scatterplot.

STEPS:

There are several steps to perform the data analysis.

- Understanding the data:

We will first look up the dataset before starting the project. The dataset was cleaned but it contained huge data .so we need to sort out meaningful data as per requirement.

- Discussing the problem statement:

After analyzing the dataset we have discussed the problem statement. After this we decided our task and initialized with it. The problem statement were based on variable we have taken for analysis.

- Data cleaning:

The next task is data cleaning which is easy with this dataset. As mentioned in above we have different column such as year ,day, month ,city etc. we have integer , float, string as a data-type.

- Exploratory Data Analysis:

After cleaning the data we have to target some column for data analysis. Matching the data with correct suitable problem by python libraries to result some visualization. These gives us information from various charts and graphs.

- Data visualization:

The EDA part make more clear about data in a picture and graphical form. We perform matplotlib and seaborn libraries of python for the data analysis. Using this we have plotted bar graph and scatterplot .

- **DATA ANALYSIS BY TEAM MEMBERS**

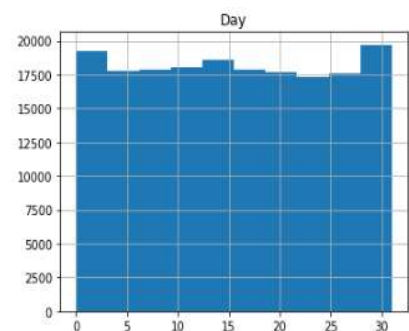
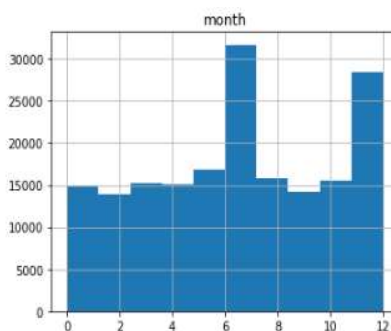
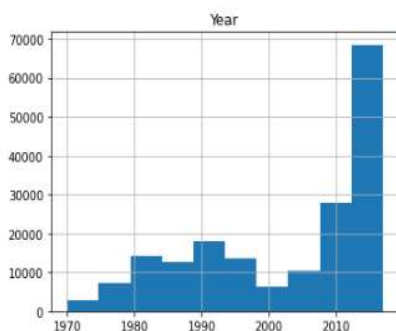
Analysis by Minal Kharbade

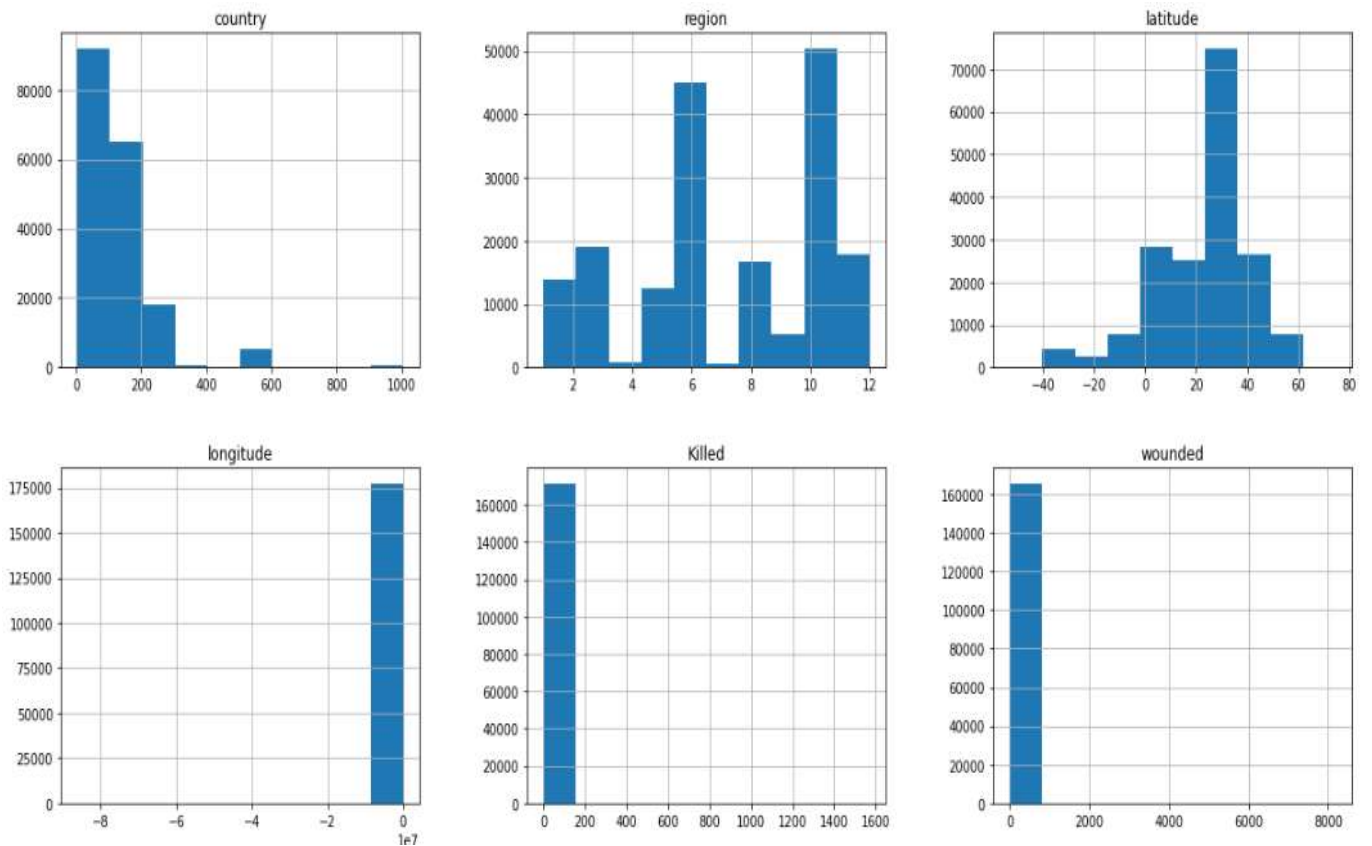
General analysis

Data Description

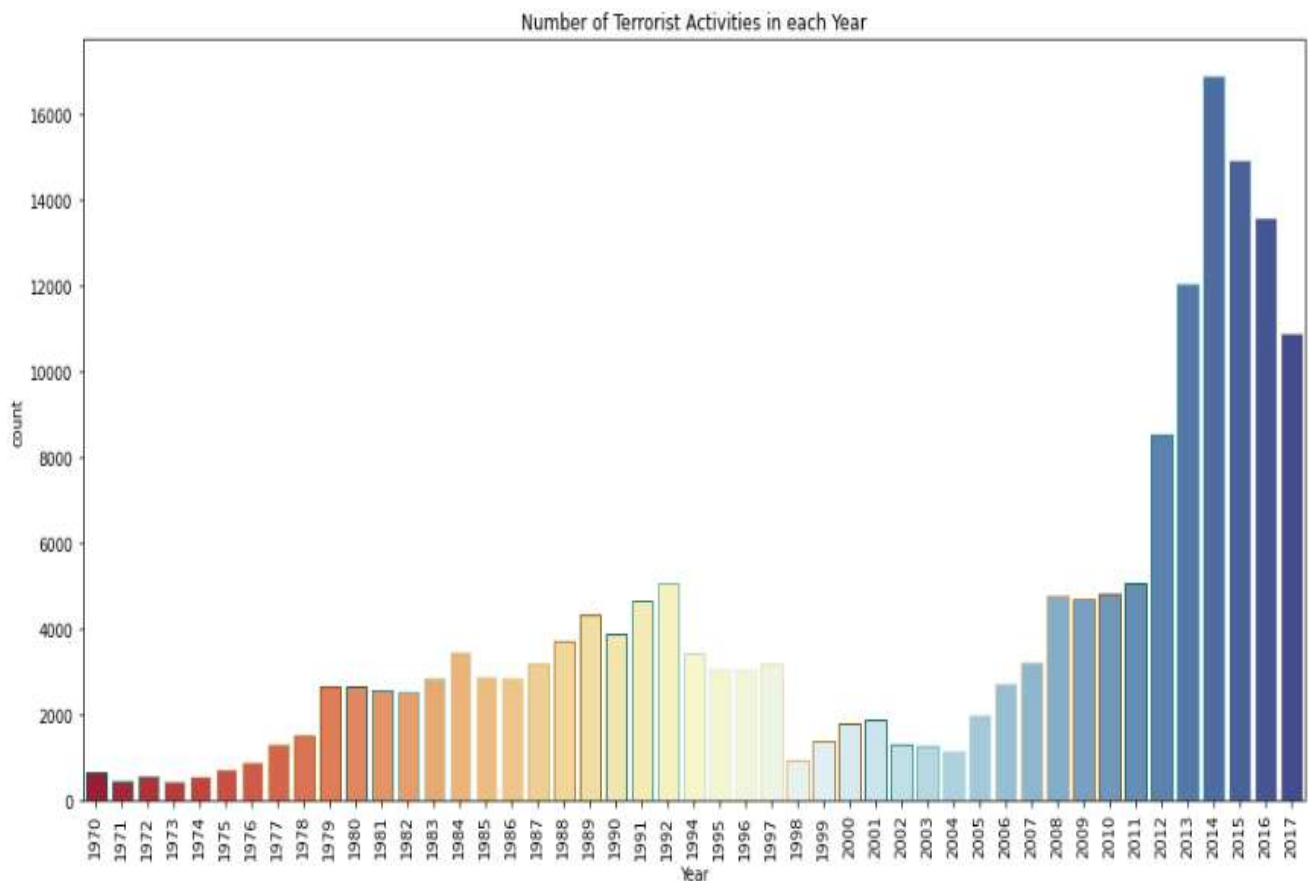
- By using the head() method we explored the the rows
- By using figsize function we have specify the width and height of a figure in unit.
- By using hist function we have plotted the histogram.
- By using the describe ()function we got the numerical values of the data frames.

Distribution of data on each series in the data frame





No of terrorist activities in year



Since 2012 there has large no of terrorist activities.

Analysis by Deveshya Gupta

Analysis based on sorted dataset

Drop the missing values

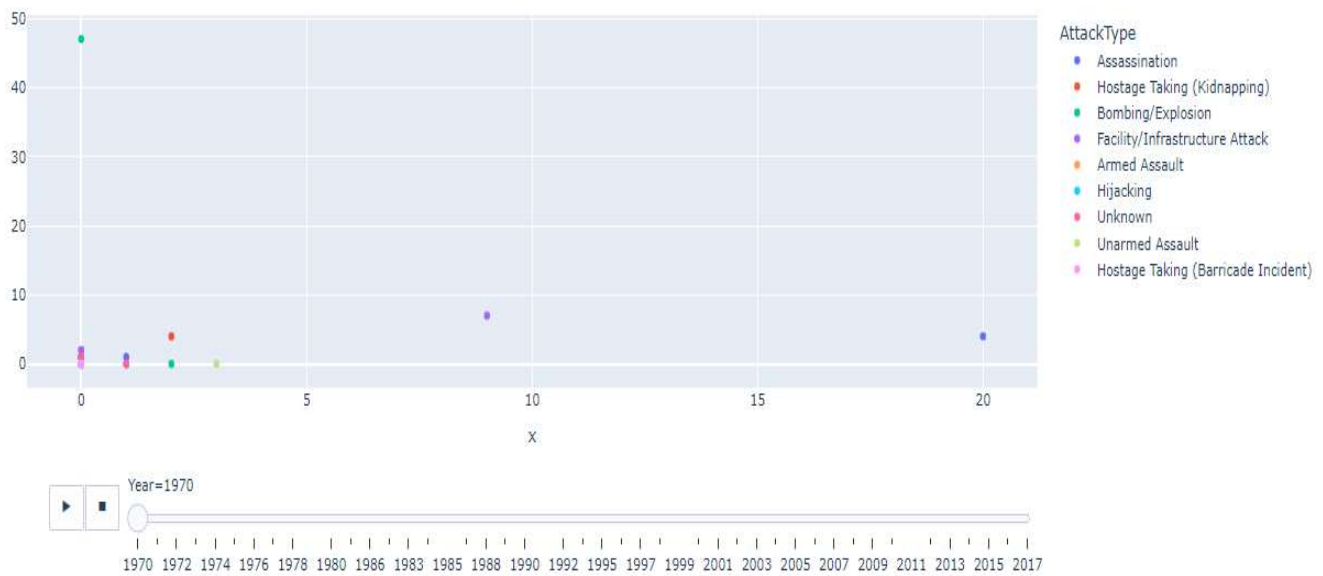
- Using heat-map function we have shown correlation between different variable.
- With scatterplot each data point is represented as a marked point, whose location is given by x and y columns.
- Using bar-plot function we have shown bargraph of different terrorist organization.

Heat-map showing correlation between different variable

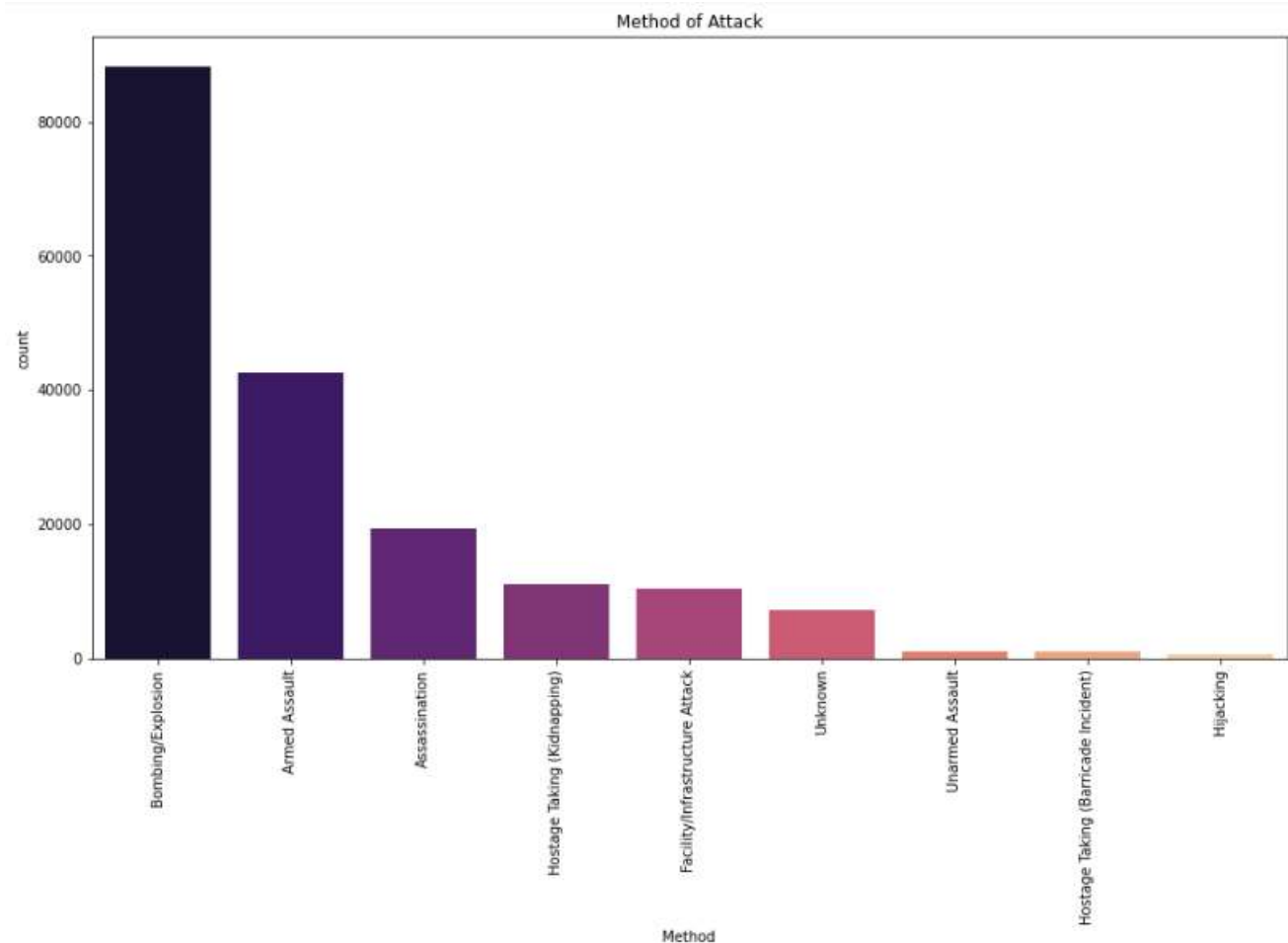


Number of casualties vs. killed people in each country for each year

Number of Casualties vs Killed people in each country for each year

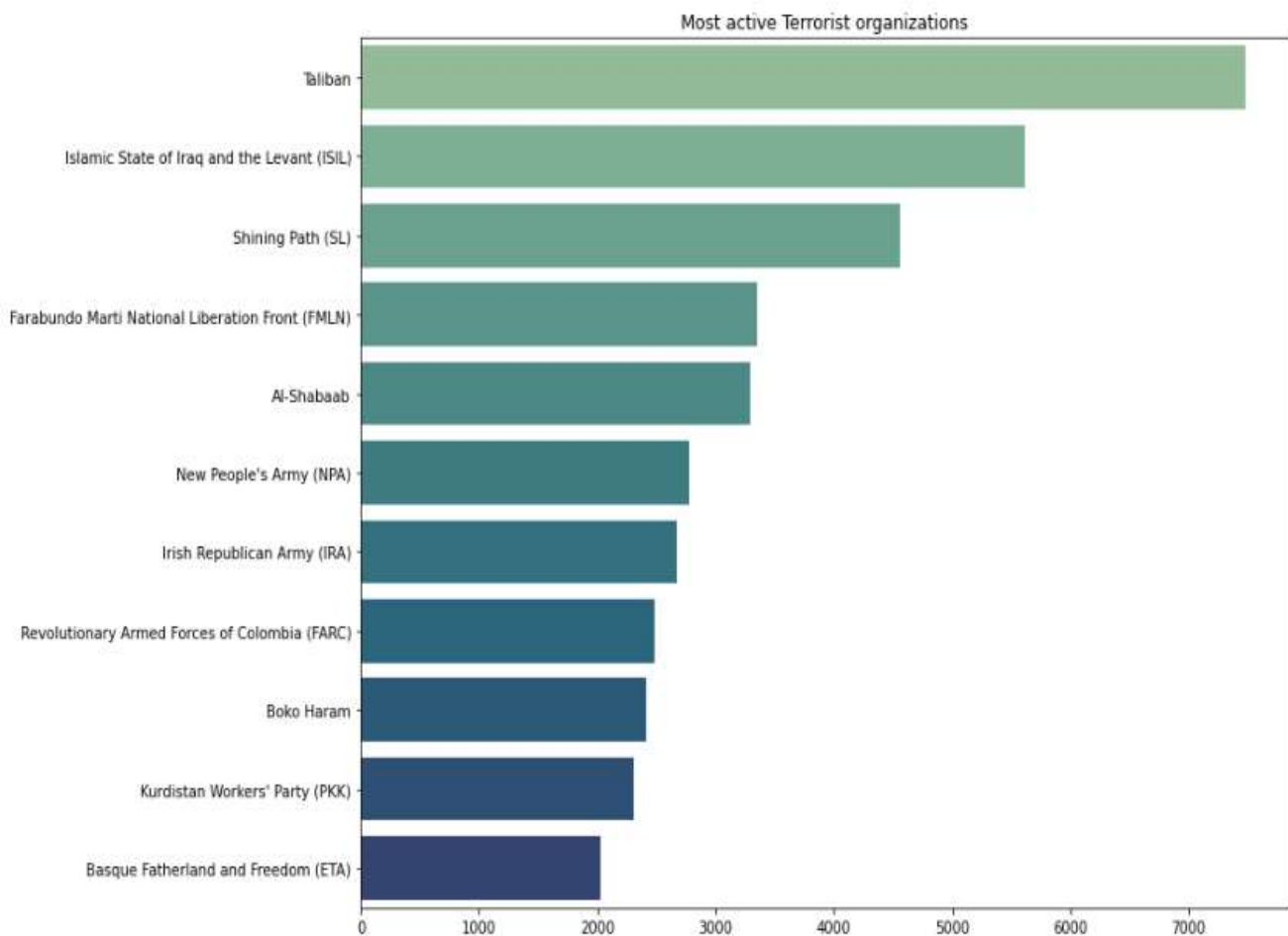


Various methods of attack used in terrorism



Different types of method use for attack such as bombing, assassination, hostage taking etc. Bombing method is mostly used for attack.

Terrorist organization have carried out their operation in country



Above bar graph shows various organization carried out their activities .so here we can say Taliban is the most active organization .

CONCLUSION:

- Throughout the project we learned the things from problem statement to understand the view of analyzing Global Terrorism.
- We have compared various parameters during the EDA (exploratory data analysis) and studied its behavior.
- From the overall analysis we conclude that since 2012 number of terrorist activities increased.
- Mostly bombing/explosion are widely used for attacks.
- Most targeted area for attack are private citizens and property.
- Among all terrorist organization Taliban is the most active organization.